J Arid Land
https://doi.org/10.1007/s40333-018-0064-4

Homogeneity analysis of streamflow records in arid and semi-arid regions of northwestern Iran

Majid KAZEMZADEH, Arash MALEKIAN*

University of Tehran, Karaj 31587-77871, Iran

Abstract: Homogeneity analysis of the streamflow time series is essential to hydrological modeling, water resources management and climate change studies. In this study, five absolute homogeneity tests and one clustering approach were used to determine the homogeneity status of the streamflow time series (over the period 1960–2010) in 14 hydrometric stations of three important basins (i.e., Aras River Basin, Urmia Lake Basin and Sefid-Roud Basin) in northwestern Iran. Results of the Buishand range test, von Neumann ratio test, cumulative deviation test, standard normal homogeneity test and Pettitt test for monthly streamflow time series detected that about 42.26%, 38.09%, 33.33%, 39.28% and 68.45% of the streamflow time series were inhomogeneous at the 0.01 significance level, respectively. Streamflow time series of the stations located in the eastern parts of the study area or within the Urmia Lake Basin were mostly homogeneous. In contrast, streamflow time series in the stations of the Aras River Basin and Sefied-Roud Basin showed inhomogeneity at annual scales. Based on the overall classification for the monthly and annual streamflow series, we determined that about 45.60%, 11.53% and 42.85% of the time series were categorized into the 'useful', 'doubtful' and 'suspect' classes according to the five absolute homogeneity tests. We also found the homogeneity patterns of the streamflow time series by using the clustering approach. The results suggested the effectiveness of the clustering approach for homogeneity analysis of the streamflow time series in addition to the absolute homogeneity tests. Moreover, results of the absolute homogeneity tests and clustering approach indicated obvious decreasing change points of the streamflow time series in the 1990s over the three basins, which were mostly related to the hydrological droughts.

Keywords: streamflow time series; homogeneity test; clustering analysis; inhomogeneity; Urmia Lake; northwestern Iran

1 Introduction

The hydrological time series increasingly show non-stationarity behavior and their variables such as streamflow and precipitation series do not show a consistent mean or median for a long period of time mainly due to the natural and anthropogenic changes (Rougé et al., 2013). Homogeneity analysis in the hydrological time series used for water resources management and project planning is important for detecting the accuracy and validity of data. Meanwhile, homogeneity analysis of the streamflow time series is essential to hydrological modeling and climate change studies. Hydrologic time series sequences are the results of particular natural conditions and might be shown irregular fluctuations when natural conditions of the river basin are relatively steady. However, the

*Corresponding author: Arash MALEKIAN (E-mail: Malekian@ut.ac.ir; Mkazemzadeh@ut.ac.ir)

Received 2017-08-12; revised 2018-02-03; accepted 2018-02-13

hydrologic time series sequences can exhibit evident trends or jumps, if natural conditions change noticeably (Wong et al., 2006). In this regard, for detecting the same patterns of the time series, two important methods can be used to analyze the homogeneity of the streamflow time series, including statistical homogeneity test and clustering approach. Statistical homogeneity test has been used in hydrological analysis for the detection of non-homogeneity which suggested by Hirsch et al. (1982), Hirsch and Slack (1984), Hirsch (1988) and McCuen (2002). Generally, homogeneities of the streamflow time series can be determined by two different statistical approaches, including relative and absolute homogeneity tests (Peterson et al., 1998). Relative homogeneity tests are few to be recommended to use, considering the neighboring stations are hypothetically homogeneous. However, absolute homogeneity tests are recommended to use if the two time series of the neighboring stations are not adequately correlated (Wijngaard et al., 2003).

Several researches have applied the statistical homogeneity tests and clustering approach for the hydro-climatic time series (e.g., Conrad and Pollak, 1950; Wijngaard et al., 2003; Kahya et al., 2008; Sahin and Cigizoglu, 2010; Dikbas et al., 2013; Hosseinzadeh Talaee et al., 2014; Seyam and Othman, 2015; Omar et al., 2017). For examples, Wijngaard et al. (2003) used the statistical homogeneity tests to analyze the climatic variables in the European continent and reported that about 25% of the precipitation series and 94% of the temperature series were labelled 'suspect' or 'doubtful' over the period 1901–1999. Sahin and Cigizoglu (2010) applied four homogeneity tests (i.e., standard normal homogeneity test, Pettitt test, Von Neumann ratio test and bivariate test) for the meteorological time series of Turkey, and revealed that the four homogeneity tests showed the same results of the time series in most of the cases. Dikbas et al. (2013) utilized the K-means clustering method to classify the maximum annual flows and identify the hydrological homogenous groups in Turkey. They suggested that the homogenous regions defined by the K-means clustering method can be used for regional flood frequency analysis. Seyam and Othman (2015) analyzed the long-term variation of annual streamflow regime in the Selangor River over a 50-year period (from 1961 to 2010) by using normality and homogeneity tests (including Shapiro-Wilk test and Pettitt test), and they found that the maximum annual streamflow totally increased whereas the minimum annual streamflow significantly decreased with respect to time.

Understanding the characteristics and sensitivities of the alternative tests is very important to analyze the homogeneity of the streamflow time series, mainly due to the large array of available statistical homogeneity tests. Applying a statistical test which is insensitive to a specific type of homogeneity can result in a failure to determine the homogeneity (McCuen, 2002). In this study, we analyzed the homogeneity of the streamflow time series (over the period 1960–2010) in 14 hydrometric stations of three important basins in northwestern Iran using five absolute homogeneity tests which are mostly used in the hydro-climatic data and clustering analysis.

2 Study area and data collection

The study area is located in northwestern Iran, which consists of three important basins, namely, Aras River Basin (ARB), Urmia Lake Basin (ULB) and Sefid-Roud Basin (SRB). The ARB covers an area of 41.0×10^3 km^2. The second one, ULB, covers an area of 51.8×10^3 km^2 and includes three important rivers, i.e., Aji Chai River, Zarrineh-Roud River and Simineh-Roud River. The SRB with an area of 60.5×10^3 km^2 is located between the Zagros Mountain Ranges and Alborz Mountain Ranges. The study area shows different soil and vegetation types. In this study, 14 available hydrometric stations with valid and adequate data over 1960–2010 were selected. Spatial distribution of these 14 stations is shown in Figure 1, and characteristics of the hydrometric stations are shown in Table 1.

3 Methods

3.1 Cumulative deviations

The adjusted partial sums or cumulative deviations (S_k^*) from the mean can be used for testing the homogeneity of the data series (Eq. 1).

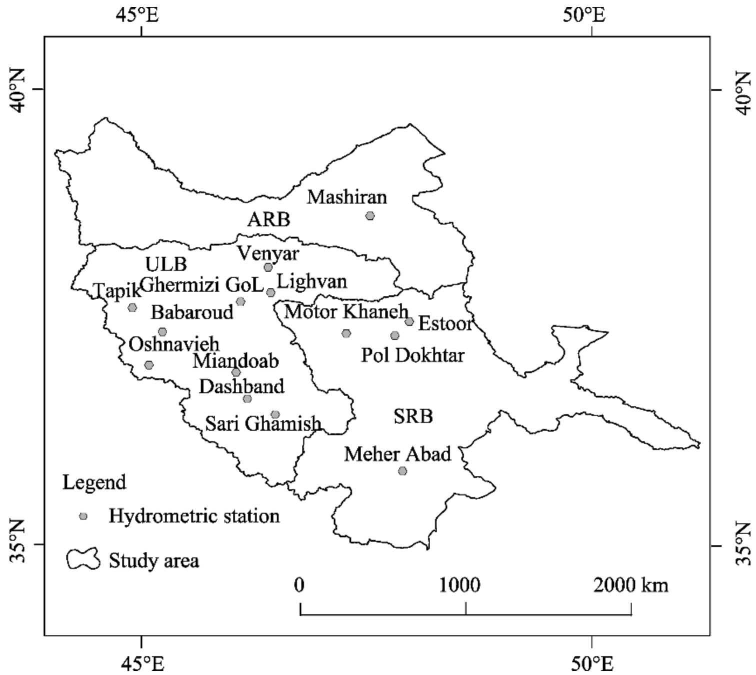

Fig. 1 Spatial distribution of the 14 hydrometric stations in the study area. ARB, Aras River Basin; ULB, Urmia Lake Basin; SRB, Sefid-Roud Basin.

Table 1 Details of the 14 selected hydrometric stations

No.	Station	River	Latitude	Longitude	Elevation (m a.s.l.)
1	Babaroud	Barandooz Chai	37°24'00"N	45°14'00"E	1297
2	Dashband	Simineh Roud	36°39'00"N	46°10'00"E	1350
3	Estoor	Ghezel-Ozan	37°30'46"N	47°58'19"E	930
4	Ghermizi Gol	Gambar Chai	37°44'00"N	46°06'00"E	1800
5	Lighvan	Lighvan	37°50'11"N	46°26'08"E	2150
6	Mashiran	Darreh Roud	38°41'10"N	47°32'01"E	705
7	Meher Abad	Talvar	35°51'15"N	47°53'50"E	1650
8	Miandoab	Simineh Roud	36°57'00"N	46°03'00"E	1310
9	Motor Khaneh	Aydogmosh	37°22'57"N	47°16'20"E	1050
10	Oshnavieh	Gelaz Chai	37°02'00"N	45°05'00"E	1480
11	Pol Dokhtar	Ghezel-Ozan	37°21'30"N	47°48'42"E	1100
12	Sari Ghamish	Sari Ghamish	36°29'00"N	46°29'00"E	1380
13	Tapik	Nazlou Chai	37°40'00"N	44°54'00"E	1399
14	Venyar	Aji Chai	38°07'55"N	46°24'22"E	1450

$$S_k^* = \sum_{i=1}^{k}\left(Y_i - \overline{Y}\right) \quad (k = 1, 2, \ldots, n), \tag{1}$$

where Y_i is the observed value of the climate parameter i; $\overline{Y}$ is the sample mean; and n is the number of records in the data series. The rescaled adjusted partial sums (S_k^{**}) can be obtained by using Equation 2:

$$S_k^{**} = \frac{S_k^*}{D_x} \quad (k = 1, 2, \ldots, n), \tag{2}$$

where D_x is the sample standard deviation, which can be calculated as follows:

$$D_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2. \tag{3}$$

In this study, we used Q statistic to measure the sensitivity to departures from homogeneity:

$$Q = \max\left|S_k^{**}\right| \ (0 \le k \le n). \tag{4}$$

The higher values of Q statistic indicate non-homogeneity in the time series. The critical values of Q for some specified values of n are given by Buishand (1982), which were based on the 19,999 synthetic sequences of Gaussian random numbers. The critical values of Q statistic in the cumulative deviation test are shown in Table 2.

Table 2 The 1% and 5% critical values of $Q/\sqrt{n}$ statistic in the cumulative deviation test and $R/\sqrt{n}$ statistic in the Buishand range test as function of n (Buishand, 1982)

n	Critical value of $Q/\sqrt{n}$ statistic		Critical value of $R/\sqrt{n}$ statistic	
	95%	99%	95%	99%
10	1.14	1.29	1.28	1.38
20	1.22	1.42	1.43	1.60
30	1.24	1.46	1.50	1.70
40	1.26	1.50	1.53	1.74
50	1.27	1.52	1.55	1.78
100	1.29	1.55	1.62	1.86
∞	1.36	1.63	1.76	2.00

3.2 Absolute homogeneity tests

3.2.1 Buishand range test

R is another statistic which can be used for homogeneity analysis (Eq. 5):

$$R = \max S_k^{**} - \min S_k^{**} \ (0 \le k \le n). \tag{5}$$

The critical values of R statistic in the Buishand range test are shown in Table 2.

3.2.2 Von Neumann ratio test

The ratio of the mean square successive difference to the variance is defined as the von Neumann ratio (N), which is described by Buishand (1982) and can be calculated as follows:

$$N = \sum_{i-1}^{n-1}\frac{(Y_i - Y_{i+1})^2}{\sum\limits_{i-1}^{n}(Y_i - \bar{Y})^2}. \tag{6}$$

The value of N statistic tends to be lower than the expected value in cases that the sample contains a break, and the values of N statistic are ≥2 if the sample has rapid variations in the mean (Bingham and Nelson, 1981). However, this test cannot identify the year of break. The critical values of N statistic in the von Neumann ratio test are given in Table 3.

Table 3 The 1% and 5% critical values of N statistic in the von Neumann ratio test as function of n (Buishand, 1982)

	Critical value of N statistic					
	n=20	n=30	n=40	n=50	n=70	n=100
99%	1.04	1.20	1.29	1.36	1.45	1.54
95%	1.30	1.42	1.49	1.54	1.61	1.67

3.2.3 Standard normal homogeneity test

The standard normal homogeneity test usually shows higher sensitivity to breaks near the first and the end parts of the time series. Alexandersson (1986) has defined the $T(k)$ statistic for the standard normal homogeneity test:

$$T(k) = k\bar{z}_1^{-2} + (n-k)\bar{z}_2^{-2} \quad (k=1,2,\ldots,n), \tag{7}$$

$$\bar{z}_1 = \frac{1}{k}\sum_{i=1}^{k}\left(Y_i - \bar{Y}\right)/s \ \text{ and } \ \bar{z}_2 = \frac{1}{k-1}\sum_{i=k+1}^{k}\left(Y_i - \bar{Y}\right)/s, \tag{8}$$

where $\bar{z}_1$ and $\bar{z}_2$ are the parameters of $T(k)$ statistic; k is the years of record; $\bar{Y}$ is the mean of time series; Y_i is the annual series which will be tested; and s is the standard deviation.

Based on calculation of the mean of the first k years and the last n–k years of the time series, we assumed that the $T(k)$ statistic reaches its maximum value when a break occurs at the year k. The T_0 statistic in the standard normal homogeneity test is described as follows (Eq. 9):

$$T_0 = \max T_k \quad (1 \le k \le n), \tag{9}$$

The critical values of T_0 statistic are given in Table 4. According to the standard normal homogeneity test, the null hypothesis will be rejected if T_0 statistic exceeds the critical values.

Table 4 The 1% critical values of T_0 statistic in the standard normal homogeneity test as function of n (Jarušková, 1996) and the 5% critical values of T_0 statistic in the standard normal homogeneity test as function of n (Alexandersson and Moberg, 1997)

	Critical value of T_0 statistic					
	n=20	n=30	n=40	n=50	n=70	n=100
99%	9.56	10.45	11.01	11.38	11.89	12.32
95%	6.95	7.65	8.10	8.45	8.80	9.15

3.2.4 Pettitt test

Pettitt test as a non-parametric approach is suitable to detect the breaks which occur near the middle of the time series. This approach is based on the Wilcoxon test developed by Pettitt (1979). The ranks r_1, r_2, …, r_k of the Y_1, Y_2, …, Y_k were used to calculated the X_k statistic in the Pettitt test:

$$X_k = \sum_{i=1}^{k} r_i - k(n+1) \quad (k=1,2,\ldots,n). \tag{10}$$

Based on the Pettitt test, the absolute value of X_k statistic will reach to its maximum value if a break occurs in a given year (Eq. 11).

$$X_k = \max|X_k| \quad (1 \le k \le n). \tag{11}$$

The critical values of X_k statistic in the Pettitt test suggested by Pettitt (1979) are presented in Table 5.

Table 5 The 1% and 5% critical values of X_k statistic in the Pettitt test as function of n (Pettitt, 1979)

	Critical value of X_k statistic					
	n=20	n=30	n=40	n=50	n=70	n=100
99%	71	133	208	293	488	841
95%	57	107	167	235	393	677

3.2.5 Clustering approach of the streamflow time series

The clustering approach is able to classify several time series in different clusters. Many researchers have applied different kinds of classification (Lagacherie et al., 1997; Ramachandra Rao and Srinivas, 2006; Kahya et al., 2008; Hsu and Li, 2010; Zahraie and Roozbahani, 2011; Dikbas et al., 2013; Kousari et al., 2013). The clustering approach was applied to facilitate the homogeneity analysis in addition to the absolute homogeneity tests in this study. The time series usually show relatively different averages and variances, then it is essential to normalize the initial time series before the clustering (Kousari et al., 2013). In this study, the hierarchical clustering was considered and the Ward's algorithm and squared Euclidean distance were preferred. The hierarchical clustering approach aims to group a set of cases in such a way that cases in the same group or

cluster are more similar to each other, which leads to a minimized variance within a group or cluster (Everitt, 1993; Kahya et al., 2008). We classified the normalized streamflow time series based on the dendrograms. We also used the overall classification and qualitative interpolation of the five absolute homogeneity tests. The classification is based on the number of the homogeneity tests which reject the null hypothesis (Wijngaard et al., 2003).

3.3 Data analysis

The Microsoft Office, XLSTAT and ArcGIS were used for data analysis and figure mapping.

4 Results and discussion

The descriptive statistics of the annual streamflow series shown in Table 6 can better reflect the streamflow regime patterns in the study area. The results indicated that in the SRB, Estoor station with an average annual discharge of 79.43 m^3/s had the highest annual water yield while Lighvan station with an average annual discharge of 0.78 m^3/s showed the lowest annual water yield. Moreover, Ghermizi Gol station with coefficient of variation (CV) of 98.80% exhibited the highest temporal variability, while Lighvan station with the CV value of 28.86% had the lowest temporal variability.

Table 6 Descriptive statistics of the annual streamflow series for different hydrometric stations

Station	Annual streamflow				
	Mean (m^3/s)	Maximum (m^3/s)	Minimum (m^3/s)	Standard deviation (m^3/s)	CV (%)
Babaroud	8.20	19.27	1.09	3.94	48.07
Dashband	15.18	39.43	2.25	8.71	57.39
Estoor	79.43	241.03	7.16	45.70	57.53
Ghermizi Gol	1.14	8.65	0.31	1.13	98.80
Lighvan	0.78	1.39	0.36	0.23	28.86
Mashiran	15.23	34.40	2.76	7.06	46.35
Meher Abad	7.44	22.10	1.15	4.67	62.79
Miandoab	16.33	45.96	2.38	9.57	58.60
Motor Khaneh	5.07	13.85	0.16	3.29	64.88
Oshnavieh	1.41	3.26	0.16	0.68	47.87
Pol Dokhtar	48.38	119.31	5.10	28.24	58.38
Sari Ghamish	52.22	146.96	10.46	25.81	49.42
Tapik	12.29	28.70	3.95	5.87	47.74
Venyar	12.13	42.84	0.71	7.63	62.89

Note: CV, coefficient of variation.

In this study, five absolute homogeneity tests were used and their results for the cumulative deviation test and Buishand range test are shown in Tables 7 and 8, respectively. The observation data of each station were analyzed for a significance level of 0.01, then inhomogeneities were detected. For the monthly streamflow series, the results of $Q/\sqrt{n}$ statistic in the cumulative deviation test showed that 42.26% of the streamflow time series of the stations were inhomogeneous. The monthly streamflow series in the Lighvan station were found homogenous in all 12 months while the monthly streamflow series in the Mashiran station in 9 months out of 12 months were detected with high inhomogeneity. Also, the results of Buishand range test indicated that about 38.09% of the streamflow time series were inhomogeneous at the 0.01 significance level. Furthermore, in most of the streamflow time series, the cumulative deviation test detected the change year in the 1990s at the 0.01 significance level. According to the study of Hawkins (1997), the cumulative deviation test is more sensitive when a break happens in the middle part of a data series.

Table 7 Results of $Q/\sqrt{n}$ statistic in the cumulative deviation test for different hydrometric stations

Station	$Q/\sqrt{n}$ statistic											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Babaroud	1.91**	2.09**	1.19	1.27	0.90	1.50	1.64**	1.22	1.62**	1.50	1.10	1.75**
Dashband	0.97	1.31	0.97	1.61**	1.67**	1.60**	1.92**	1.98**	1.94**	2.02**	0.61	1.49
Estoor	1.57**	1.70**	1.22	1.66**	1.82**	1.53**	1.28	1.28	1.46	1.81**	1.81**	1.58**
Ghermizi Gol	1.24	1.06	0.72	0.98	0.98	1.17	1.23	0.71	1.28	1.84**	1.63**	1.00
Lighvan	0.76	1.09	0.92	1.01	1.09	0.89	0.92	0.92	0.90	0.80	1.04	1.31
Mashiran	2.14**	2.32**	2.38**	2.35**	1.93**	1.81**	1.56**	0.55	1.33	1.03	1.62**	1.96**
Meher Abad	1.80**	1.71**	1.41	1.46	1.42	0.95	0.99	1.04	1.42	2.12**	1.96**	1.09
Miandoab	1.03	0.72	0.57	1.63**	1.82**	1.37	1.21	1.35	1.31	1.08	0.73	1.52**
Motor Khaneh	1.04	1.22	0.93	1.35	1.92**	1.60**	1.23	1.53**	1.70**	1.36	1.89**	1.21
Oshnavieh	1.39	1.59**	1.03	1.51	1.11	1.38	0.87	1.12	1.68**	1.72**	0.99	1.00
Pol Dokhtar	1.60**	1.66**	1.36	1.66**	1.75**	1.44	1.07	1.18	1.37	1.76**	1.78**	1.82**
Sari Ghamish	1.42	1.63**	1.20	0.78	1.19	0.79	2.23**	2.43**	2.14**	1.64**	1.34	1.47
Tapik	1.60**	1.60**	0.63	1.01	1.17	1.47	1.33	1.16	1.24	1.59**	1.58**	1.68**
Venyar	1.70**	1.70**	1.55**	1.73**	1.49	1.70**	1.14	0.70	1.11	1.14	1.60**	1.97**

Note: ** indicates the significance at the 0.01 level.

Table 8 Results of $R/\sqrt{n}$ statistic in the Buishand range test for different hydrometric stations

Station	$R/\sqrt{n}$ statistic											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Babaroud	2.10**	2.21**	1.28	1.68	1.24	1.50	1.64	1.22	1.62	1.74	1.39	1.97**
Dashband	1.53	1.92**	1.08	1.82**	1.67	1.60	1.92**	1.98**	1.94**	2.21**	1.19	2.28**
Estoor	2.17**	2.26**	1.64	2.08**	1.82**	1.53	1.57	1.29	1.60	2.13**	2.11**	2.13**
Ghermizi Gol	1.24	1.09	1.27	0.98	1.01	1.17	1.23	0.95	1.31	1.84**	1.63	1.00
Lighvan	1.01	1.27	1.00	1.09	1.14	1.01	0.98	0.92	1.40	1.07	1.40	1.54
Mashiran	2.14**	2.32**	2.38**	2.35**	1.93**	1.81	1.56	0.84	1.34	1.07	1.62	1.96**
Meher Abad	2.46**	2.23**	1.99**	2.09**	1.66	1.19	1.58	1.50	1.88**	2.84**	2.45**	1.55
Miandoab	1.41	1.36	0.82	1.82**	1.86**	1.46	2.19**	1.87**	2.32**	1.87**	0.93	2.35**
Motor Khaneh	1.77	1.97**	1.30	1.64	1.95**	1.64	1.32	1.53	1.70	1.44	1.92**	2.17**
Oshnavieh	2.13**	2.10**	1.28	1.67	1.19	1.38	1.05	1.12	1.68	1.72	1.10	1.76
Pol Dokhtar	2.28**	2.18**	1.88**	2.16**	1.76	1.54	1.44	1.25	1.70	2.12**	2.10**	2.51**
Sari Ghamish	1.72	1.63	1.20	0.97	1.46	1.43	2.23**	2.43**	2.36**	3.20**	2.31**	1.94**
Tapik	1.83**	1.76	0.99	1.39	1.44	1.49	1.39	1.16	1.24	1.79**	1.87**	1.96**
Venyar	1.88**	1.90**	1.75	1.82**	1.53	1.70	1.16	1.10	1.13	1.25	1.60	2.15**

Note: ** indicates the significance at the 0.01 level.

The outputs of three absolute homogeneity tests, namely, von Neumann ratio test, standard normal homogeneity test and Pettitt test, are presented in Tables 9, 10 and 11, respectively. The von Neumann ratio test, standard normal homogeneity test and Pettitt test identified that about 33.33%, 39.28% and 68.45% of the streamflow time series were considered as inhomogenous at monthly scales at the 0.01 significant level, respectively. The results of von Neumann ratio test showed that the total number of inhomogeneity of the streamflow time series in clod seasons (autumn and winter) was more than that in warm seasons (spring and summer). The highest number of inhomogeneity in monthly streamflow series was found in February based on the von Neumann ratio test. Homogeneity analysis of precipitation series in Iran indicated that the total number of inhomogeneity was more in clod season months or winter months than in warm season months (Hosseinzadeh Talaee et al., 2014). It was previously reported that in the semi-arid regions, the variability of the streamflow time series was usually larger in warm season months than in clod season months mainly due to the natural and anthropogenic changes (e.g., Hereford and Webb,

1992; Barlow et al., 2001; Das et al., 2011). In this study, the standard normal homogeneity test and Pettitt test did not show any specific particular pattern for the monthly streamflow series over the period 1960–2010. The monthly streamflow series in the Mashiran, Estoor, Pol Dokhtar and Mothor Khaneh stations exhibited inhomogeneous patterns.

Table 9 Results of *N* statistic in the von Neumann ratio test for different hydrometric stations

Station	*N* statistic											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Babaroud	0.95**	0.78**	1.36**	1.84	1.49	1.40	1.55	1.84	1.71	1.56	1.67	0.87**
Dashband	1.66	1.29**	1.89	1.90	1.64	1.61	1.50	1.43	1.37	1.28**	1.74	1.30**
Estoor	0.86**	0.98**	1.57	1.78	1.49	1.58	1.90	1.61	1.38	1.10**	1.04**	0.80**
Ghermizi Gol	1.87	1.27**	2.35	1.99	2.13	1.84	1.96	1.85	1.66	0.83**	0.82**	1.82
Lighvan	2.08	1.82	1.96	1.87	2.02	1.84	1.95	1.68	1.56	1.81	1.48	1.83
Mashiran	0.96**	0.82**	0.93**	1.25**	1.65	1.73	2.04	2.09	1.84	1.60	1.70	1.15**
Meher Abad	0.52**	0.59**	1.57	1.87	1.66	1.75	1.30**	1.41	1.41	0.66**	0.82**	1.41
Miandoab	1.52	1.49	1.63	1.85	1.54	1.54	1.07**	1.35**	1.08**	1.12**	2.17	1.22**
Motor Khaneh	1.22**	0.83**	2.00	1.76	1.35**	1.60	1.97	1.79	0.88**	1.83	1.14**	1.00**
Oshnavieh	1.42	1.36**	1.67	1.93	1.79	1.62	1.88	1.93	1.33**	1.83	2.03	1.50
Pol Dokhtar	0.71**	0.80**	1.38	1.80	1.43	1.43	1.79	1.66	1.37	1.02**	1.05**	0.48**
Sari Ghamish	1.40	1.44	1.50	1.92	1.85	1.74	0.68**	0.46**	0.52**	0.32**	1.35**	1.47
Tapik	0.81**	1.12**	1.80	1.38	1.37	1.39	1.53	1.52	0.97**	0.58**	0.92**	0.74**
Venyar	1.29**	1.48	1.40	1.59	1.71	1.58	2.20	2.12	2.11	1.56	1.76	0.96**

Note: ** indicates the significance at the 0.01 level.

Table 10 Results of T_0 statistic in the standard normal homogeneity test for different hydrometric stations

Station	T_0 statistic											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Babaroud	17.73**	21.57**	6.62	7.48	3.96	10.92	13.07**	6.37	13.29**	6.03	10.44	14.91**
Dashband	4.34	6.79	3.85	11.39**	11.86**	10.88	17.60**	15.40**	15.91**	1.52	19.04**	9.96
Estoor	12.56**	14.94**	7.58	14.02**	16.13**	10.92	7.95	8.16	9.90	17.15**	15.91**	12.80**
Ghermizi Gol	9.80	5.81	5.16	7.00	6.98	9.94	10.98	7.26	10.21	29.36**	24.07**	13.00**
Lighvan	4.18	7.24	14.69**	11.86**	8.11	5.00	4.59	6.86	4.14	4.21	3.92	7.35
Mashiran	20.58**	26.13**	26.65**	25.25**	16.30**	12.97**	9.73	1.88	8.91	11.89**	4.18	17.24**
Meher Abad	16.49**	14.83**	9.60	10.41	9.76	4.07	4.77	5.21	9.81	19.66**	23.26**	6.00
Miandoab	4.78	4.21	1.75	12.06**	14.14**	7.98	7.40	8.91	8.29	2.53	5.40	10.37
Motor Khaneh	4.91	6.91	3.92	8.47	18.57**	13.11**	7.37	15.45**	19.23**	18.13**	9.16	6.90
Oshnavieh	8.24	10.79	4.21	9.20	5.10	7.49	3.85	8.39	15.02**	4.03	12.52**	4.47
Pol Dokhtar	13.00**	14.07**	9.46	14.00**	15.25**	10.47	5.85	6.48	8.70	16.49**	16.28**	17.26**
Sari Ghamish	9.08	11.97**	10.46	3.13	7.22	4.54	29.09**	36.17**	30.34**	7.51	16.57**	9.77
Tapik	12.44**	12.46**	1.76	4.77	6.54	10.16	8.66	6.51	7.59	12.15**	12.95**	14.09**
Venyar	13.47**	11.44**	10.17	13.73**	10.19	11.44**	5.22	2.47	6.24	15.04**	9.08	17.58**

Note: ** indicates the significance at the 0.01 level.

Results of the five absolute homogeneity tests for annual streamflow series are shown in Figure 2. The cumulative deviation test and Buishand range test determined that streamflow time series in 5 and 6 stations (out of 14 stations) showed homogenous, respectively. Meanwhile, the results of standard normal homogeneity test were the same with those of cumulative deviation test. Furthermore, the von Neumann ratio test and Pettitt test identified streamflow time series in 5 and 3 stations (out of 14 stations) as homogeneous, respectively. Based on the above-mentioned results, we inferred that the Pettitt test is more sensitive than the other absolute homogeneity tests in the

Table 11 Results of X_k statistic in the Pettitt test for different hydrometric stations

Station	X_k statistic											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Babaroud	410**	423**	221	297**	152	324**	426**	300**	375**	385**	273	362**
Estoor	429**	379**	357**	391**	417**	388**	381**	394**	447	446**	448**	441**
Dashband	326**	317**	263	387**	380**	399**	462**	530**	469	491**	266	368**
Ghermizi Gol	305**	207	127	157	144	285**	273	233	280	396**	282	276
Lighvan	108	214	139	153	241	158	272	250	233**	211	208	270
Mashiran	504**	486**	471**	491**	409**	405**	320**	174	365**	260	370**	452**
Meher Abad	441**	417**	396**	394**	374**	316**	195	212	410**	461**	467**	457**
Miandoab	317**	203	224	383**	412**	340**	263	314**	280**	298**	286	384**
Motor Khaneh	301**	343**	364**	367**	419**	387**	289**	383**	329**	425**	414**	337**
Oshnavieh	298**	348**	279	330**	222	314**	230	255	395**	454**	166	240
Pol Dokhtar	411**	405**	369**	401**	429**	411**	309**	337**	408**	450**	440**	423**
Sari Ghamish	386**	354**	265	193	275	190	407**	400**	386**	366**	296**	378**
Tapik	344**	342**	136	257	256	332**	349**	254	319**	377**	324**	382**
Venyar	429**	371**	396**	438**	345**	313**	237	190	255	327**	453**	456**

Note: ** indicates the significance at the 0.01 level.

determination of homogeneity of the streamflow time series. As a good consequence, the results of the five homogeneity tests determined that streamflow time series in the stations located in the east of the study area or within the ULB were mostly found to be homogeneous. In contrast, the annual streamflow series of all stations located in the ARB and SRB were inhomogeneous. The streamflow time series in the three stations of the ULB namely Ghermizi Gol, Lighvan and Sari Ghamish stations were identified as homogeneous by all homogeneity tests. Generally, since the absolute homogeneity tests have different sensitivities to the variability and changes of the data series in a given station, then the outputs of the tests sometimes have discrepancies. These discrepancies were also suggested by Wijngaard et al. (2003), Feng et al. (2004), Sahin and Cigizoglu (2010) and Hosseinzadeh Talaee et al. (2014).

The overall classification determined that about 45.60% of the streamflow time series were labeled as 'useful', which suggested no obvious evidence of inhomogeneity in the streamflow time series. According to the study of Wijngaard et al. (2003), the 'useful' class is adequately homogeneous for variability and trend analyses. About 11.53% of the streamflow time series were labelled as 'doubtful'. In other words, 11.53% of the monthly and annual streamflow series in the stations were critical in terms of the possible inhomogeneities. The third class known as 'suspect' refers to the situation that the null hypothesis is rejected by three or four tests at the 0.01 significance level. Finally, 42.85% of the streamflow time series were labeled in this class. It is obvious that the streamflow time series labeled as 'suspect' class have insufficient credibility and should not be applied in issues which are related to time series studies.

Figure 3 indicated the different clusters of the normalized annual streamflow series with some clear fluctuations in the clusters. Moreover, downtrend alterations or change points in the streamflow time series occurred after 1990 for all stations, and the results from the five absolute homogeneity tests were almost the same. Meanwhile, fluctuation of the third cluster (Fig. 3c) was more dramatic than those of the first (Fig. 3a) and second (Fig. 3b) clusters. It was obvious that the first cluster (Fig. 3a) contained most of the streamflow time series compared with the other two clusters (Figs. 3b and c). Since these curves did not show a spatial distribution in the study area, then the distributions of various clusters of specific streamflow time series were mapped in Figure 4. The results showed that most of the streamflow time series were in the first cluster. For the monthly streamflow series, we found that only Estoor station was in the third cluster in January, April, May, November and December. As mentioned earlier, the number of inhomogeneity of the streamflow time series was more in the third cluster than in the first and second clusters. The results indicated that the clustering time series could identify the homogeneity of the streamflow time

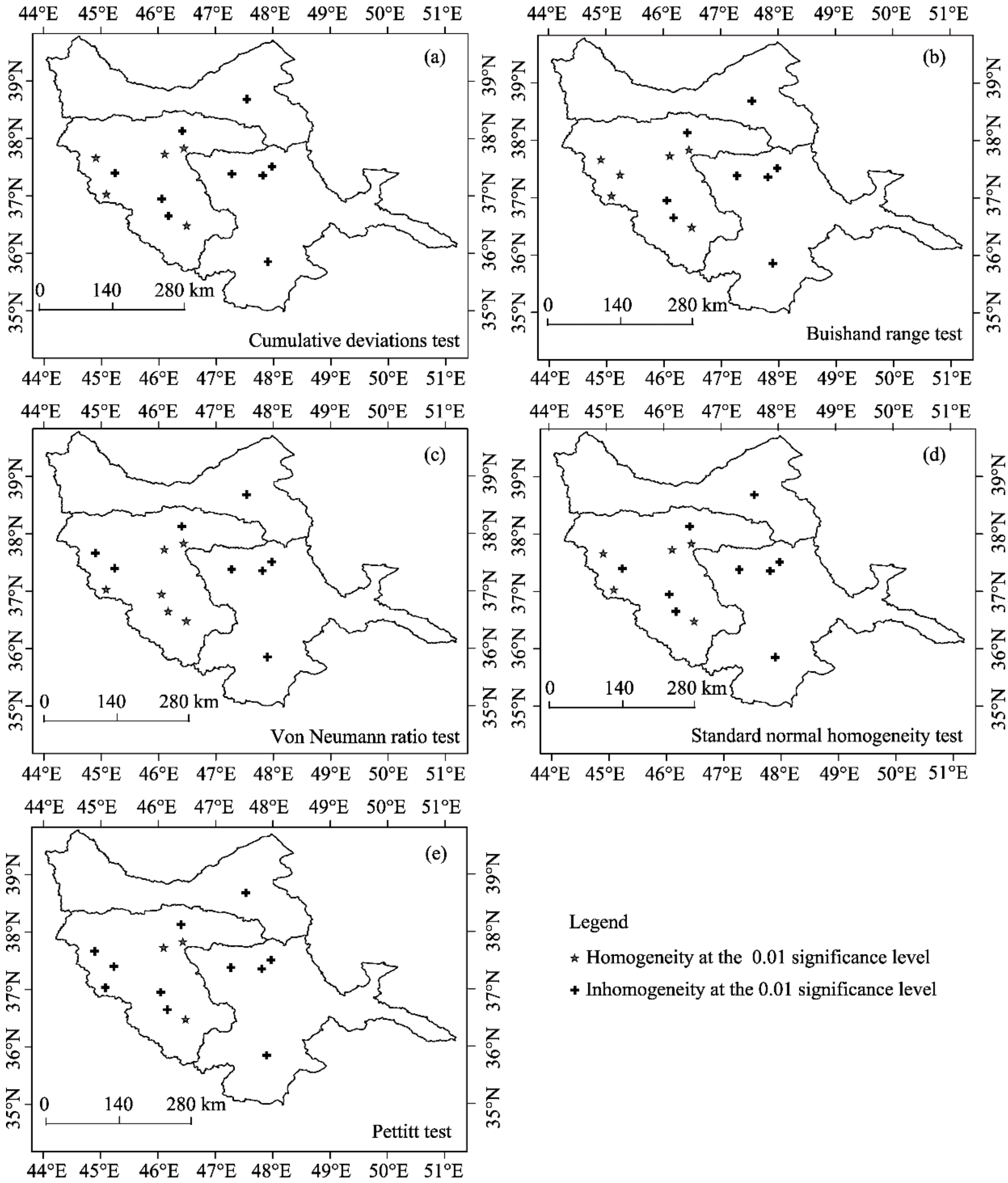

Fig. 2 Outputs of five homogeneity tests for annual streamflow series in the 14 hydrometric stations over the period 1960–2010

series in most of the stations. Homogeneity tests determined that most of the streamflow time series in the Estoor station were inhomogeneous. Meanwhile, most of the streamflow time series in this station were classified in the third cluster which showed high fluctuations and inhomogeneity. All time series of streamflow except for February in the Estoor and Pol Dokhtar stations were classified in the second and third clusters. Furthermore, inhomogeneity was also detected in most of the streamflow time series in the stations by the five absolute homogeneity tests.

Generally, the hydrometric stations located in the SRB were classified in the second and third clusters in most streamflow time series while the stations in the ULB were categorized in the first cluster which exhibited low inhomogeneity of the streamflow time series. Therefore, the results of this study clearly proved the effectiveness of the clustering approach for homogeneity analysis of

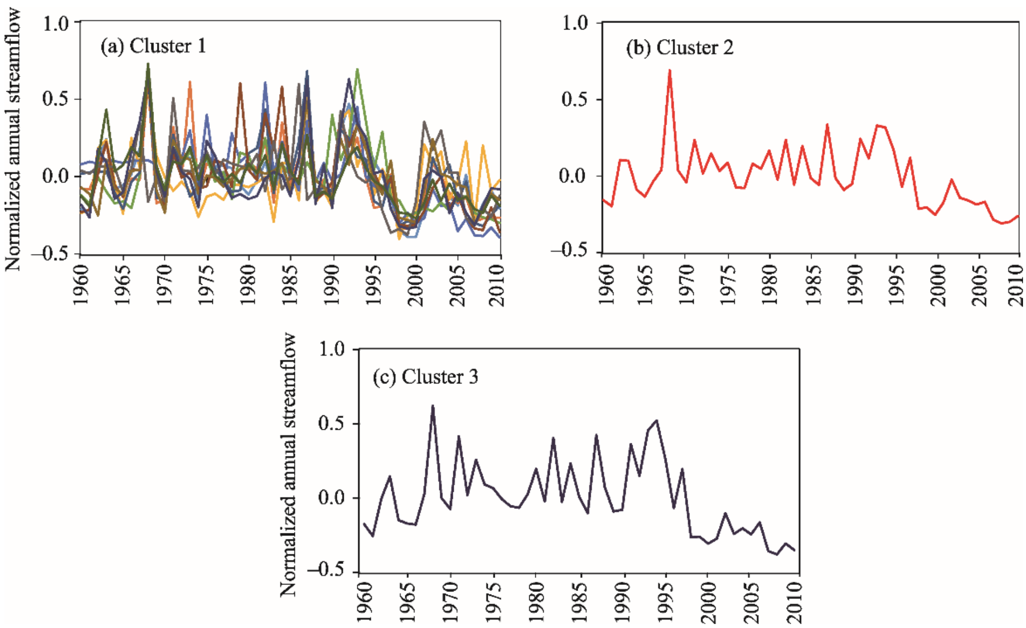

Fig. 3 Normalized annual streamflow series of cluster 1 (a), cluster 2 (b) and cluster 3 (c)

the streamflow time series in addition to the homogeneity tests. The clustering analysis to classify the streamflow time series into regions with relatively similar streamflow patterns by Kahya et al. (2008) indicated that the regions with the same streamflow patterns will not be covered enough with the climatic zones in Turkey. Kousari et al. (2013) has suggested that clustering approach is suitable to analyze the trends in the climatic variables over the Iran.

Results of the absolute homogeneity tests and clustering approach indicated that stations located in the three basins (SRB, ARB and ULB) have the same fluctuations and alterations of the streamflow time series in most of the study period. Moreover, obvious decreasing change points were observed in the 1990s over the three basins. It is clear that inhomogeneity of the streamflow time series cannot be completely explained by the local changes of the hydrometric stations or watershed conditions, such as changes in the locations of stations, land use changes, etc. From the results we concluded that alternations in the streamflow time series of the study area are mostly related to the hydrological droughts. The droughts occurred in the 1990s were more devastating to water resources and agriculture, resulting in accelerated urbanization (Agrawala et al., 2001; Yazdani and Haghsheno, 2008). According to the study of Raziei et al. (2009), over half of population in Iran has been affected by prolonged droughts in the 1990s. Nikbakht et al. (2013) and Tabari et al. (2013) analyzed the streamflow droughts in northwestern Iran and found the most severe streamflow scarcities in the 1990s. The decreasing trend of the streamflow time series in the study area can be considered as an important reason of inhomogeneities in most of the stations.

5 Conclusions

The current research considered five absolute homogeneity tests and clustering approach for determining the homogeneity of the streamflow time series (over the period 1960–2010) in three important basins (SRB, ARB and ULB) located in northwestern Iran. The results of cumulative deviation test, Buishand range test, von Neumann ratio test, standard normal homogeneity test and Pettitt test for the monthly streamflow series showed that about 42.26%, 38.09%, 33.33%, 39.28% and 68.45% of the time series were inhomogeneous at the 0.01 significance level, respectively. Among the five absolute homogeneity tests, Pettitt test was found to be much more sensitive than the other four homogeneity tests in determination of homogeneity of the streamflow time series. Results of the five absolute homogeneity tests determined that the streamflow time series in the

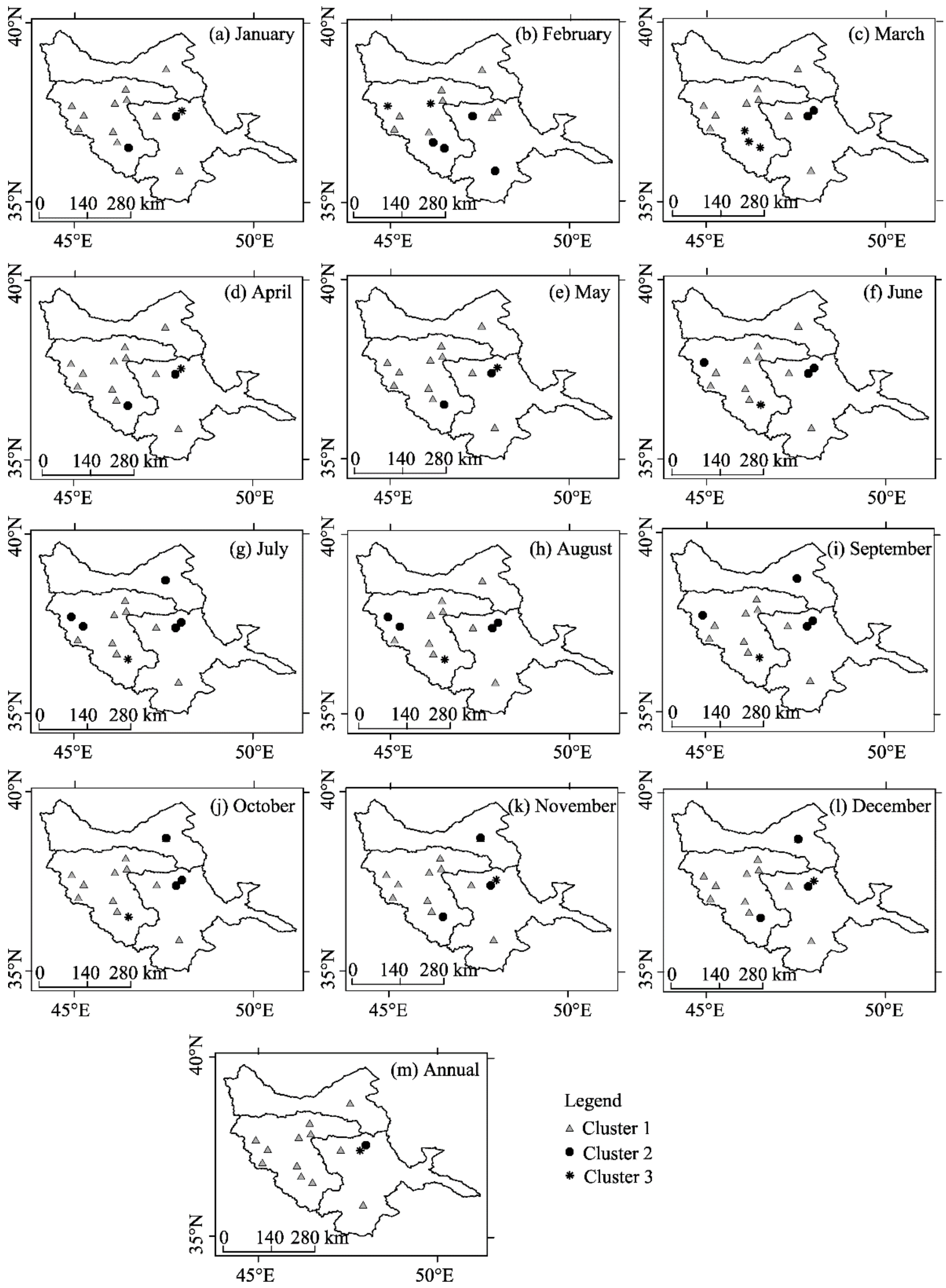

Fig. 4 Spatial distributions of three clusters in monthly and annual streamflow series

stations located in the east of the study area or within the ULB were mostly found to be homogeneous. In contrast, the streamflow time series in all stations of the ARB and SRB were inhomogeneous at annual scales. Since the absolute homogeneity tests have various sensitivities to the variability and changes in the streamflow time series of the station, the outputs of these tests may show discrepancies in some cases. The overall classification and qualitative interpolation of the monthly and annual streamflow series based on the five absolute homogeneity tests showed

that about 45.60%, 11.53% and 42.85% of the series were categorized into 'useful', 'doubtful' and 'suspect' classes, respectively. Furthermore, the effectiveness of the clustering approach for analyzing the homogeneity of the streamflow time series was proved in addition to the absolute homogeneity tests. Both the results of absolute homogeneity tests and clustering approach indicated relatively similar fluctuations and alterations in the streamflow time series of the stations. Moreover, obvious decreasing change points of the streamflow time series were detected in the 1990s over the three basins. It can be inferred that inhomogeneity and alternations in the streamflow time series of the three basins are related to the natural and anthropogenic conditions of the basins.

References

Agrawala S, Barlow M, Cullen H, et al. 2001. The drought and humanitarian crisis in Central and Southwest Asia: a climate perspective. Palisades, N.Y.: International Research Institute for Climate and Society, 50–160.

Alexandersson H. 1986. A homogeneity test applied to precipitation data. Journal of Climateology, 6(6): 661–675.

Alexandersson H, Moberg A. 1997. Homogenization of Swedish temperature data. Part 1: homogeneity test for linear trends. International Journal of Climatology, 17: 25–34.

Barlow M, Nigam S, Berbery E H. 2001. ENSO, Pacific decadal variability, and US summertime precipitation, drought, and stream flow. Journal of Climate, 14(9): 2105–2128.

Bingham C, Nelson L S. 1981. An approximation for the distribution of the Von Neumann ratio. Technometrics, 23(3): 285–288.

Buishand T A. 1982. Some methods for testing the homogeneity of rainfall records. Journal of Hydrology, 58(1–2): 11–27.

Conrad V, Pollak L W. 1950. Methods in Climatology. Cambridge, Massachusetts, USA: Harvard University Press, 1–459.

Das T, Pierce D W, Cayan D R, et al. 2011. The importance of warm season warming to western U.S. streamflow changes. Geophysical Research Letters, 38(23): L23403.

Dikbas F, Firat M, Koc A C, et al. 2013. Defining homogeneous regions for streamflow processes in Turkey using a K-Means clustering method. Arabian Journal for Science and Engineering, 38(6): 1313–1319.

Everitt B. 1993. Cluster Analysis. New York: Halsted Press, 1–130.

Feng S, Hu Q, Qian W. 2004. Quality control of daily meteorological data in China, 1951–2000: A new dataset. International Journal of Climatology, 24(7): 853–87.

Hawkins M. 1997. Testing a sequence of observations for a shift in location. Journal of the American Statistical Association, 72(357): 180–186.

Hereford R, Webb R H. 1992. Historic variation of warm-season rainfall, Southern Colorado Plateau, southwestern U.S.A. Climatic Change, 22(3): 239–256.

Hirsch R M, Slack J R, Smith R A. 1982. Techniques of trend analysis for monthly water quality data. Water Resources Research, 18(1): 107–121.

Hirsch R M, Slack J R. 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resources Research, 20(6): 727–732.

Hirsch R M. 1988. Statistical methods and sampling design for estimating step trends in surface-water quality. Water Resources Bulletin, 24(3): 493–503.

Hosseinzadeh Talaee P H, Kouchakzadeh M, Shifteh Some'e B S. 2014. Homogeneity analysis of precipitation series in Iran. Theoretical and Applied Climatology, 118(1–2): 297–305.

Hsu K C, Li S T. 2010. Clustering spatial–temporal precipitation data using wavelet transform and self-organizing map neural network. Advances in Water Resources, 33(2): 190–200.

Jarušková D. 1996. Change–point detection in meteorological measurement. Monthly Weather Review, 124: 1535–1543.

Kahya E, Demirel M C, Bég O A. 2008. Hydrologic homogeneous regions using monthly streamflow in Turkey. Earth Sciences Research Journal, 12(2): 181–193.

Kousari M R, Ahani H, Hendi-zadeh R. 2013. Temporal and spatial trend detection of maximum air temperature in Iran during 1960–2005. Global and Planetary Change, 111: 97–110.

Lagacherie P, Cazemier D R, van Gaans P F, et al. 1997. Fuzzy k-means clustering of fields in an elementary catchment and extrapolation to a larger area. Geoderma, 77(2–4): 197–216.

McCuen R H. 2002. Modeling Hydrologic Change: Statistical Methods. Boca Raton, Florida: CRC Press , 1–433.

Nikbakht J, Tabari H, Hosseinzadeh Talaee P. 2013. Streamflow drought severity analysis by percent of normal index (PNI) in northwest Iran. Theoretical and Applied Climatology, 112(3–4): 565–573.

Omar M A, Bağçac C, Şarlak N. 2017. Homogeneity Analysis of Precipitation Series in North Iraq. IOSR Journal of Applied Geology and Geophysics, 5(3): 57–63.

Peterson T C, Vose R, Schmoyer R, et al. 1998. Global Historical Climatology Network (GHCN) quality control of monthly temperature data. International Journal of Climatology, 18: 1169–1179.

Pettitt A N. 1979. A non-parametric approach to the change-point problem. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(2): 126–135.

Ramachandra Rao A, Srinivas V V. 2006. Regionalization of watersheds by hybrid-cluster analysis. Journal of Hydrology, 318(1–4): 37–56.

Raziei T, Saghafian B, Paulo A A, et al. 2009. Spatial patterns and temporal variability of drought in western Iran. Water Resources Management, 23(3): 439–455.

Rougé C, Ge Y, Cai X. 2013. Detecting gradual and abrupt changes in hydrological records. Advances in Water Resources, 53: 33–44.

Sahin S, Cigizoglu H K. 2010. Homogeneity analysis of Turkish meteorological data set. Hydrological Processes, 24(8): 981–992.

Seyam M, Othman F. 2015. Long-term variation analysis of a tropical river's annual streamflow regime over a 50-year period. Theoretical and Applied Climatology, 121(1–2): 71–85.

Tabari H, Nikbakht J, Hosseinzadeh Talaee P. 2013. Hydrological drought assessment in Northwestern Iran based on streamflow drought index (SDI). Water Resources Management, 27(1): 137–151.

Wijngaard J B, Klein Tank A M G, Können G P. 2003. Homogeneity of 20th century European daily temperature and precipitation series. International Journal of Climatology, 23(6): 679–692.

Wong H, Hu B Q, Ip W C, et al. 2006. Change-point analysis of hydrological time series using grey relational method. Journal of Hydrology, 324(1–4): 323–338.

Yazdani S, Haghsheno M. 2008. Drought management and recommended solutions on how to deal with drought. American-Eurasian Journal of Agricultural & Environmental Sciences, 2(Suppl. 1): 64–68.

Zahraie B, Roozbahani A. 2011. SST clustering for winter precipitation prediction in southeast of Iran: comparison between modified K-means and genetic algorithm-based clustering methods. Expert Systems with Applications, 38(5): 5919–5929.